DOCUMENT RESUME

ED 360 545 CE 064 330

AUTHOR	Hall, William; Saunders, John
TITLE	Getting to Grips with Assessment.
INSTITUTION	National Centre for Vocational Education Research, Leabrook (Australia).
REPORT NO	ISBN-0-86397-325-4
PUB DATE	[93]
NOTE	51p.
AVAILABLE FROM	National Centre for Vocational Education Research Ltd., 252 Kensington Road, Leabrook, South Australia 5068, Australia.
PUB TYPE	Guides - Classroom Use - Teaching Guides (For Teacher) (052)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Annotated Bibliographies; *Competency Based Education; Criterion Referenced Tests; *Evaluation Methods; Foreign Countries; Models; Prior Learning; Program Development; Self Evaluation (Individuals); *Student Evaluation; Test Bias; Testing; *Testing Programs; Test Reliability; *Test Use; Test Validity; *Vocational Education

ABSTRACT

This booklet has been written to help persons interested in assessment of education and training programs in general and competency-based vocational education programs in particular. The following topics are covered in the individual sections: the meaning of the term "assessment"; the importance of assessment; curriculum models; percentages and comparison; sites where assessment should be conducted; assessment methods; appropriate persons for conducting assessments; valid, reliable, and economical assessment; bias in assessment; steps involved in planning assessment; recognition of prior learning; and problems and pitfalls of assessment results. The final section is an annotated bibliography of 23 publications concerning assessment in competency-based vocational education programs. (MN)

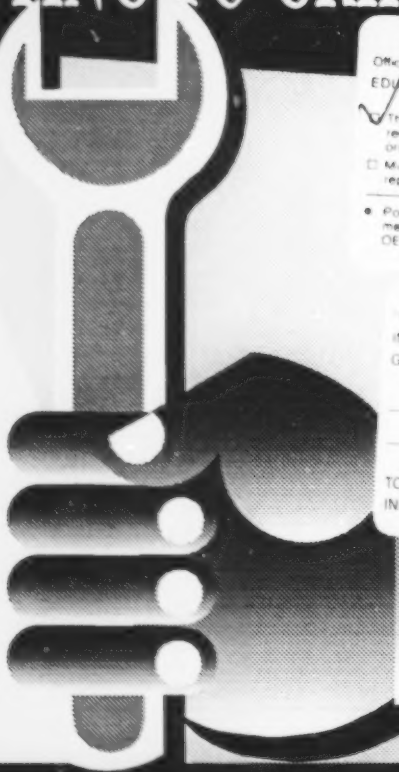
****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 360 545

NCVER

THE NATIONAL CENTRE FOR
VOCATIONAL EDUCATION RESEARCH

GETTING TO GRIPS WITH



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

M Dews

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

ASSESSMENT

CE064330

BEST COPY AVAILABLE

GETTING TO GRIPS WITH ASSESSMENT

William Hall and John Saunders

3

ABOUT THE SERIES

The "Getting to Grips With . . ." series has been written for the general reader who wants to understand important trends in vocational education and training.

Each booklet is in two parts. The first part comprises a description of the subject matter in a manner which is intended to be clear to any interested layperson. The second part gives an annotated list of publications for those who want to read further.

Other titles in the series include: *Getting to grips with competency-based training and assessment, Getting to grips with skills audits, Getting to grips with industry restructuring, Getting to grips with key technologies, Getting to grips with on-the-job training, Getting to grips with on-the-job competencies, Getting to grips with quality, and Getting to grips with implementing CBT.*

The series editor is Dr William Hall, executive director, National Centre for Vocational Education Research Ltd. The series designer is Bain Middleton. Photographs were provided by CALS Media Unit, Adelaide College of TAFE, Robin Cornwell, Segments Photo Library and SA DETAFE.

© **National Centre for Vocational Education Research Ltd**
A.C.N. 007 967 311

ISBN: 0 86397 326 4
TD/TNC: 34.05
Published by:
National Centre for Vocational Education Research Ltd
252 Kensington Road LEABROOK SA 5068

CONTENTS

INTRODUCTION	1
WHAT IS MEANT BY ASSESSMENT?	2
WHY ASSESSMENT MATTERS	4
CURRICULUM MODEL	8
PERCENTAGES AND COMPARISONS	10
WHERE SHOULD I ASSESS?	13
HOW CAN I ASSESS?	16
WHO SHOULD CARRY OUT THE ASSESSMENT?	22
VALID, RELIABLE AND ECONOMICAL ASSESSMENT	23
BIAS IN ASSESSMENT	28
PLANNING ASSESSMENT	31
RECOGNISING PRIOR LEARNING	31
ASSESSMENT RESULTS — SOME PROBLEMS AND PITFALLS	34
AND LASTLY	38
FURTHER READING	39

INTRODUCTION

Assessment has, once again, become a big issue. We say 'once again' because interest in assessment has been cyclical. During the 1940s interest in assessment was high, especially in the development of (so called) objective tests which were used in the US armed forces. The behavioural objectives period, but especially the 1960s and 1970s, saw a massive resurgence of interest in assessment.

Today's interest has been generated by a significant interest within the vocational education sector and industry in competency-based learning. If we are to have competency standards, ways of assessing those standards are needed. Fortunately, most of the assessment techniques developed over the past fifty years are as applicable today as they were yesterday. It is not surprising, therefore, that large quantities of useful reference books are available.

Many of the books in the further reading section have been drawn upon. The workshop manual *Assessing competence in the motor industry* has also been extensively used and some of that publication's overhead transparencies incorporated in the text.

The intention of this booklet is to introduce the topic of 'assessment' to those who know little about it. The further reading section may then be used to take any part of the topic further.

WHAT IS MEANT BY ASSESSMENT?

In education, assessment means finding out how successfully someone has achieved an aim.

Assessment has two components:

- collecting information about how a person performs
- making judgements based on the information collected.

Everyone has been assessed at some time. Here are three common examples:

- a driving test
- a school examination
- operating a video-cassette recorder.

Three common assessments

In the case of the driving test, **you** (perhaps with advice from your instructor) decided when you were ready to be assessed. However, the form of the assessment was determined by others. In the case of the school examination, you probably had no influence at all on the assessment process. In the case of the video-casette recorder, you probably assessed your own competence in operating it, with no one else being involved in the process.

Assessment is almost always based on a sampling process. Just as we might taste one or two apples in order to judge a whole case of them, so too we test the performance of a student on a few tasks under prescribed conditions in order to judge the student's ability to perform a full range of skills in all circumstances. Thus, in the driving test, although not all of the necessary skills are tested in all possible conditions, it is assumed that, if the driving test is passed, the driver can successfully cope with those combinations of skills and conditions which were not tested.

Competency-based training has generated renewed interest in assessment. However, although competency-based assessment uses the same techniques as any other approach to assessment, there is a difference in emphasis. Competency-based assessment is more **practical** in emphasis. In other words, it is concerned with what people can **do** — how people perform in jobs as a result of training. The more traditional approaches to assessment usually place more emphasis on what people know.

Therefore competency-based assessment emphasises assessment of the application of skills and knowledge in a practical situation or **on the job,** whereas traditional assessment emphasises assessment of knowledge by means of written or oral testing in **an examination room.**

Even so, the place of assessment in the curriculum, the techniques of assessment, the uses of assessment, and the importance of 'reliability' and 'validity' (terms we shall consider later) all remain the same.

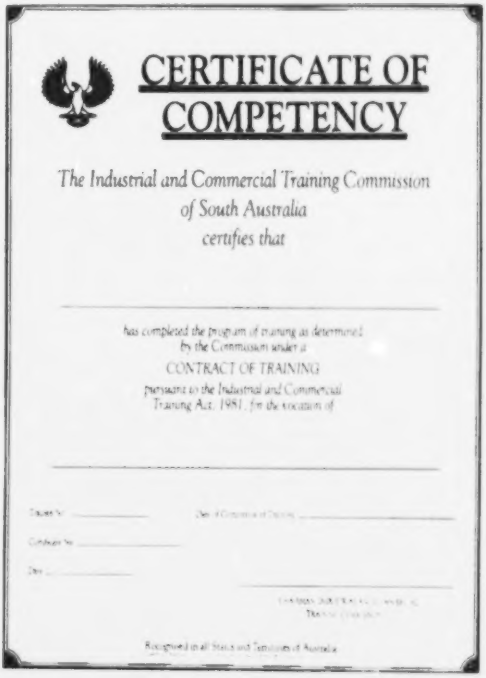
WHY ASSESSMENT MATTERS

When sitting in a doctor's waiting room, it is comforting to know that the person you will be seeing has learnt the necessary skills at a reputable institution. When we get similarly skilled people, such as electricians or plumbers, to do jobs at our homes, we need to be assured that the person doing the work is competent.

We need to be reassured that skilled people are competent

Before practising as a doctor, electrician or plumber a person must become registered with the appropriate body. That registration depends upon being successful when assessed.

Therefore one use of assessment is to determine whether or not a person has mastered a particular set of skills or body of knowledge; in this case the skills or knowledge required to perform a job. This very important type of assessment is usually referred to as assessment of **mastery** and is an essential component of competency-based training.



CERTIFICATE OF COMPETENCY

The Industrial and Commercial Training Commission of South Australia certifies that

_____

has completed the program of training as determined by the Commission under a

CONTRACT OF TRAINING

pursuant to the Industrial and Commercial Training Act, 1981, in the vocation of

_____

Proof of competence

10 BEST COPY AVAILABLE

Another use of assessment is to predict future performance; not surprisingly, referred to as **predictive** assessment. For example, entry to college or university is normally determined by assessment results gained at school.

Frequently, school-based assessment bears little relationship to what happens in tertiary education, but we persist with the myth that it does. Because society demands some sort of objective approach to college or university entry, we persist with this unsatisfactory method. That is why assessment is used as a hurdle: to trip up the required number of people and so prevent them from gaining tertiary entrance. Attempting to predict any aspect of human behaviour is always risky; so use of assessment for predictive purposes should be undertaken only with great caution. Because competency-based assessment is mainly about practical job skills, it largely avoids this trap.

Assessment is also used to promote learning. (The curriculum model which follows later illustrates this use.) In this context the assessment may take the form of **diagnostic** assessment in which learning difficulties are identified and measured to enable the most appropriate strategies to be used to assist the learner. Alternatively it may be **formative** assessment, used to provide feedback to learners during training. Such feedback enables learners to monitor their progress and make the necessary adjustments to improve their performance.

Self-assessment is a particular type of formative assessment in which the learners monitor their personal progress by assessing their own performances rather than someone else doing it for them. Self-assessment is especially suited to competency-based training programs and may be used in association with computer-managed learning and computer-assisted instruction.

Another educational application of assessment is in the measurement of **achievement**. Achievement tests enable us to determine whether a person has learned something. Such information can be important in determining whether a person possesses the prerequisite knowledge or skills necessary for entry to a course, or in determining the extent of credit which may be allowed for previous learning. This issue is taken up in more detail later in this booklet.

By looking at the performances of groups of students it is also possible to measure the effectiveness of a program of instruction or to compare one program with another. In

these cases assessment contributes evidence for the purpose of **evaluation**.

Assessment can also be used to determine **aptitudes**, **abilities** and **personal qualities** which can be helpful in providing career guidance, motivating trainees and selecting applicants for courses or employment.

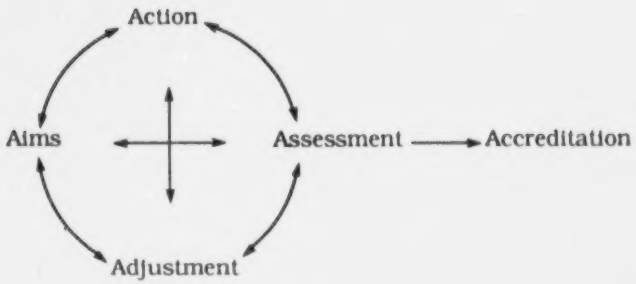
It should be noted that, in education especially, when assessment takes place will be determined by the purpose of the assessment. Assessment may occur at the beginning, during, or the end of a course.

Most importantly, assessment must never be an 'afterthought' or an irksome activity slapped on to the end of a course. Assessment must be integrated with the rest of the curriculum and must be approached as professionally as every other step in the educational process.

CURRICULUM MODEL

Assessment is an essential component of a curriculum. One simple model which illustrates this is shown on the next page. (It is based on work by Frank Halliwell.)

The model shows that **action** (training and learning) should be preceded by **aims**. These aims may be in the form of competency standards.



Assessment is needed to determine how successfully aims have been achieved. The results of assessment are sometimes used to for **accreditation**. A person may also be used as feedback to the learner so that **adjustments** may be made to learning. The process of adjustment should be continuous if assessment is used in this way.

An interesting feature of this model is that you may start curriculum planning anywhere in the cycle. For example, you may start with assessment, then devise your aims, and finally devise appropriate training.

The sad fact is that, in the past, course aims and assessment were not always clearly linked.

I learned something in school today.
I signed up for folk guitar, computer programming, stained glass art, shoemaking and a natural foods workshop.
I got spelling, history, arithmetic and two study periods.

So what did you learn?

I learned that what you sign up for
and what you get
are two different things.

Dialogue from *Peanuts*, Charles Schulz

Competency-based approaches to training allow for much closer links between what is promised and what is delivered.

PERCENTAGES AND COMPARISONS

Many of us were brought up to believe that 50% was a pass mark. We were told that above 'average' was to be strived for, and that 'below average' was failure. The difference between pass (50%) and failure (49%) could determine a person's future.

The nonsense of this may be illustrated by considering, for example, an airline pilot. What passenger would want to fly with a person who landed safely only 50% of the time? Who would want to fly with a person who got 100% for all aspects (but one) of the job? Especially if the one failure was in getting 0% for checking fuel level before take-off.

Who would want to fly with someone who consistently got 90% for landing whilst in flight simulation? Percentages,

then, mean very little. 50%, 90% or even 99% are just not good enough for pilots. Only 100% will suffice.

100% competency essential

Comparisons with others can also mean little. For example, others may get 10%, 20% 30%, 50% and 60% for flight simulation. Such percentages do not give added confidence in the trainee pilot who gets 90%, even though that person has clearly done so much better than everyone else.

Our society encourages us to compare ourselves with others. Competition is encouraged between individuals, between organisations, and between countries. It may come as a surprise, therefore, to learn that such an approach to assessment may not be helpful. (This assessment approach is called 'norm-referenced', but we won't use the jargon.)

BEST COPY AVAILABLE

It is more useful in assessment to set standards, and then to find out how successfully those standards have been achieved. (The jargon for this is 'criterion-referenced' but, once again, we won't use the term.) Competency-based assessment adopts this approach. The word 'competence' has a precise meaning when used in assessment.

Competence

- the specification of skills (including knowledge)
- the application of those skills to an agreed standard.

In developing competency-based curricula and the associated assessment, care must be taken not to remove competition entirely, because we live in a competitive world. Because of this, some people believe that grades within competencies are required, and it is possible to have grades with competency-based assessment.

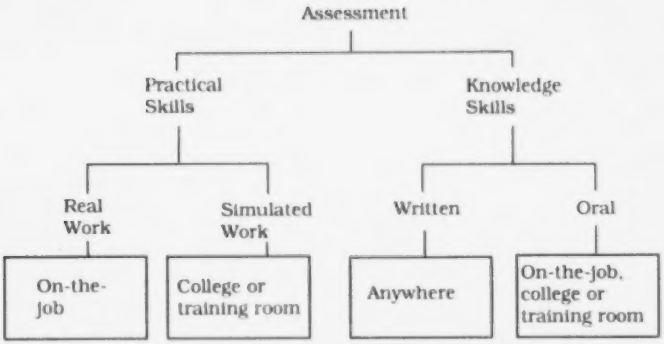
Competence and competition

WHERE SHOULD I ASSESS?

It is easy to describe many practical jobs within the home. For example, we could easily describe how to paint a wooden door (fill cracks; sand surface; apply primer; then apply undercoat and topcoat) but actually **doing** the job is rather more difficult. The same is true of the assessment of skills.

The only way to assess practical skills is by assessing practical skills

However, not all assessment needs to be done on the job, as the following diagram shows.



To explain this diagram let us consider how we might assess the competence of a maintenance fitter applying for a restricted electrical licence, which is a licence needed to disconnect and reconnect the power supply to a piece of electrical equipment. In order to perform this type of work it is obvious that the applicant is going to need both practical skills and relevant (or 'underpinning') knowledge. For example, there is not much value in being able to use a spanner correctly (practical skill) if one does not know which bolts must be loosened (knowledge), nor is there much value in knowing what voltage should exist between two terminals (knowledge) if one cannot read the voltmeter scale to check it (practical skill).

Having decided to test both practical skills and underpinning knowledge, let us look first at how and where we might assess the practical skills. Perhaps we could do it by

requiring the applicant to disconnect and reconnect a spot welder on a production line. Real work such as this provides us with a test which has high **validity** (something we shall discuss later), but there can be problems. The production line manager may want it fixed as quickly as possible, so there may not be time to allow proper observation of the process and inspection of the work done. Noise might also make communication between assessor and applicant difficult and observation of the applicant doing the job might be hampered because the equipment is not physically accessible to two people at the same time.

Faced with the above problems we may decide to switch to a simulated task in which a spot welder, or perhaps only the parts of the spot welder relevant to the task, are set up in a training area. This can provide much better conditions for assessment. However, offsetting this advantage is the fact that the applicant no longer has to contend with 'real work' conditions. One solution might be to conduct tests under both conditions: for example, disconnecting and reconnecting two items of equipment — one in the training room and one on the shop floor.

Having determined where the practical skills are to be assessed, let us now turn our attention to where we assess the knowledge.

In fact, most of the underpinning knowledge could probably be tested by the same tests as the practical skills; however to test the knowledge needed to deal with all possible contingencies would require an unacceptably large number

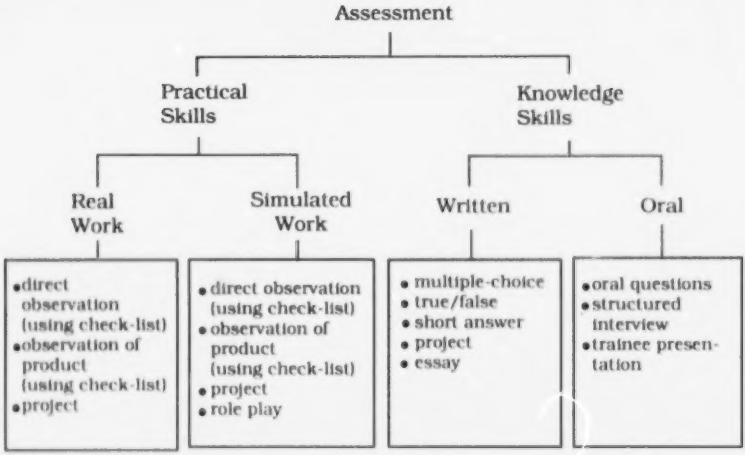
of test tasks. Instead we can assess this knowledge much more efficiently with a written or oral test. For example, instead of setting a series of practical tests to determine whether the applicant knows all of the colour codes used for electrical wiring we might instead simply ask the applicant to name all the conductor colours used in electrical wiring and state what each colour denotes. Then we only need to use practical tests to ensure the applicant knows how to apply the knowledge.

If the knowledge test is to be a written one (often referred to as a 'paper and pencil' test), it can be performed at the work site, in a classroom, in an interview room, or anywhere a satisfactory environment can be found (low noise, good lighting, comfortable temperature, etc).

Oral testing, in which the assessor asks the applicant about the task, obviously requires an environment which allows unimpeded communication, such as in an interview room or classroom.

HOW CAN I ASSESS?

There is a wide range of techniques available for assessing achievement. None is perfect, but some are more appropriate than others for particular approaches to assessment. The diagram summarises this.



To help understand this diagram let us look at how the various approaches could be applied to assessment of a student training to become a commercial artist.

One of the best ways to assess the trainee's practical skills would be **direct observation of** the trainee performing the work expected of a commercial artist; that is, observing the work **process**. So, our commercial art student could be observed preparing art work in the workplace for a genuine client (i.e. real work). If this presents any difficulties, we could instead assess the student working on an imaginary client's brief in the classroom (simulated work). By observing the student in this way we can learn a great deal about the student's level of knowledge and ability to apply it.

Sometimes it is impractical to observe a person performing a task; for reasons of limited physical access or the length of time required for observation, for example. In such cases we may be able to obtain much of the information we need by **observation or inspection of the product** of the student's work. As was the case for observation of process previously described, this may be applied to either a real or simulated work task. Returning to our example of the commercial art student — we could assess the standard of the finished art work rather than observations on the processes undertaken to produce them.

Sometimes **project work** can provide us with a useful means of assessment, especially where we wish to focus on a particular aspect of a person's knowledge or skills. As the previous diagram shows, project work is commonly used for assessing practical skills in either the workplace or a simulated work setting, and also for assessing knowledge in

written form. One of the potential weaknesses of project work is the need to establish that the completed project is the student's own work. This consideration is sometimes referred to as the **authenticity** of an assessment.

Assessment through **role play** is, by definition, an assessment of a simulated activity. It is especially useful where interpersonal skills need to be assessed. In the case of our commercial artist, role play might be used to assess the student's competence in giving a presentation to a client describing the merits of a particular poster design.

If we were to ask a group of people to give an example of assessment, most would undoubtedly cite a **written test** of knowledge. This example would be offered because written tests of knowledge are the ones we have most commonly experienced in school, in tertiary study and in our employment. Because they are so familiar we need not discuss them further in this section.

Oral assessments of knowledge are particularly useful where a need has been identified for supplementary evidence in support of another form of assessment. They are also a valuable means of assessing the knowledge of applicants who have limited English language skills or who have difficulty in reading or writing.

One of the strengths of oral forms of assessment is that they allow the assessor to persist with a line of questioning to determine whether or not a particular point is properly understood. However, this flexibility can also be a source of

weakness, for, without proper training, assessors may bias the assessment results by improperly framing questions or by leading the trainee towards a desired answer. To minimise these problems assessors must be adequately trained in techniques of oral assessment. Oral assessments can also be relatively expensive because, in most cases, they are administered individually.

Oral assessment

For applications of the various types of oral assessment listed we return to the case of our commercial art student. We might, for example, use **oral questioning** to check that the student correctly understands the complexities of a particular printing process specified in the student's poster design, or an oral test of knowledge might consist of asking the student to conduct a **presentation** to the rest of the

BEST COPY AVAILABLE

class, explaining the concepts on which a design is based. A **structured interview**, on the other hand, might form part of an assessment of an applicant's suitability for entry to a course in commercial art.

From the outline provided above it should be evident that each assessment technique has particular features which make it better suited to some applications than others. Almost always, you will find two or more techniques will need to be employed to ensure the validity and reliability of the overall result. You will need to select the approaches which are most appropriate. Numerous books are available on assessment techniques and it is suggested that you consult these when designing an assessment.

With the enormous swing in recent years to competency-based training and assessment, the importance of **observation** as an assessment technique is becoming well recognised in both industry and education. Social scientists have used this technique for many years and as a result, much is known about it.

When using the technique, remember:

- your presence will have an impact on the person being assessed; try to make that impact as slight as possible;
- questioning is a different technique from observing; observations should be done silently;
- a check-list based on aims to be achieved is essential for accurate records;
- a check-list should not emphasise the trivial;
- you won't see everything.

WHO SHOULD CARRY OUT THE ASSESSMENT?

If assessment is an integral part of a learning program, self-assessment could be appropriate. Assessment should be an integral part of the learning process. For example, projects may be both learning and assessment exercises.

However, if assessment is for certification purposes, self-assessment would be entirely inappropriate. Indeed, an independent or external assessor (someone from outside the organisation) may have to be used.

Peer group assessment can be an excellent learning opportunity. In this approach, one trainee assesses another trainee. However, it should be noted that the members of a peer group are often reluctant to judge their fellows harshly for fear that the same fate might befall them. Objective criteria can help overcome this.

WHO CAN BE AN ASSESSOR?

VEETAC (the Vocational Education Employment and Training Advisory Committee), in its recent publication *Competency Standards — Assessors* has the following to say about assessors:

> Assessors include persons working in any position involving assessment functions. They are people who are:
>
> - supervisors
> - workers on the shop floor and managers
> - persons in training or management positions
> - training provider or institutions
> - members of work teams or groups.

The point is also made that 'assessor' is not the title of a position but is the name given to a person acting in an assessment role.

VALID, RELIABLE AND ECONOMICAL ASSESSMENT

Imagine that we had to assess the data processing skills of a group of business studies trainees and the test we set them looked like this:

(1)	1 + 1 ___	(2)	2 + 1 ___	(3)	2 - 1 ___	(4)	1 + 2 ___	(5)	1 - 1 ___

Leaving aside that they might think we were crazy, it would be amazing if all trainees did not get all five calculations correct every time they attempted them, whether or not they were assessed by the same or a different person and regardless of such things as the time of day, weather conditions, or even the degree of anxiety of the trainees.

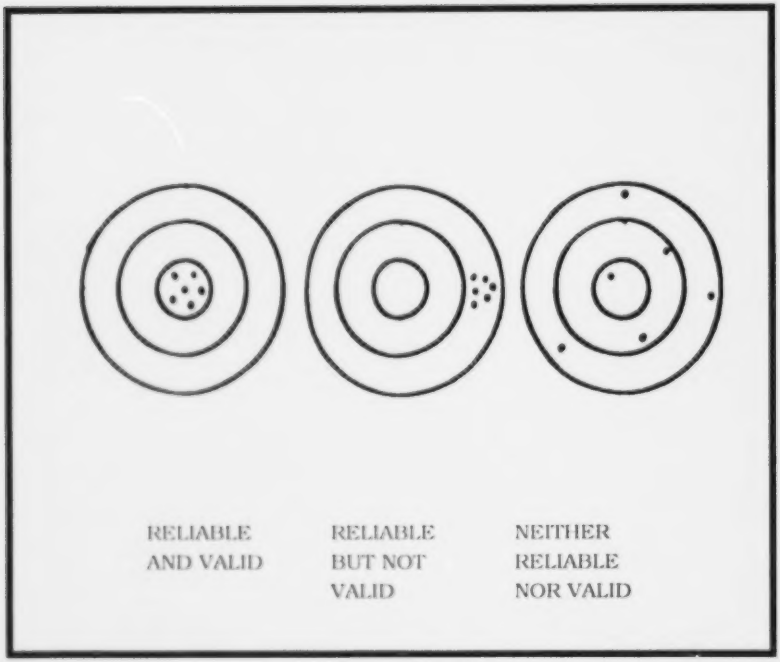
Such a test would be 100% **reliable**. However, it would be wholly **invalid** because it does not reflect what the trainees need to do in their jobs.

The terms 'reliability' and 'validity' are extremely important in assessment. Let us look at them in more detail.

Our assessment in the above example was reliable because, in normally acceptable assessment conditions it would have been almost impossible to get any other result. By **reliability** therefore, we mean how consistently an assessment can repeat the same result for each individual.

Our assessment however, was invalid because it did not test any of the skills (other than the most basic arithmetic) which the trainees needed to perform in their jobs. What should our assessor have tested? Perhaps we could have tested the person's ability to operate a computer, keyboard skills, and understanding of data base programs. By **validity** therefore, we mean how closely we achieve the aims of our assessment — are we actually measuring what we set out to measure?

The concepts of reliability and validity can be graphically illustrated by means of the following analogy to the performance of shooters aiming for the bull's eye:



Ideal assessments, those which are both reliable and valid, can be likened to the performance of rifle shooters who not only achieve consistency (**reliability**) by tightly grouping their shots, but also achieve their aim (**validity**) by placing the entire group of shots in the bull's eye.

Assessments which are reliable but invalid may be likened to the performance of accurate shooters who fail to allow for wind or incorrectly adjust their sights. The tight grouping of shots still indicates consistency (**reliability**) but failure to hit

the bull's eye means they are not longer achieving their aim (lacking in **validity**).

Assessments which lack both reliability and validity are represented by the unfortunate shooters whose shots are scattered all over the target. Not only are they failing to achieve their aim (no **validity**) but worse, they are not even shooting consistently (no **reliability**).

From this shooting analogy it may also be seen that an assessment must be reliable in order for it to be valid, but does not need to be valid in order for it to be reliable. Validity is an important issue in competency-based assessment.

One source of assessment invalidity is bias. For example, illustrations may be given or wording may be used that is biased towards a particular ethnic group, or towards one gender. Personal prejudice (whether or not you like a trainee) can also have a powerful effect on assessment validity. This is taken up in the next section.

Practical considerations are also important when assessing. There is little point in having high validity and high reliability if the testing procedure is prohibitively expensive.

COST-EFFECTIVENESS: WHICH ASSESSMENT WOULD YOU CHOOSE?

As manager of a trucking company you wish to hire a new driver for one of the company's latest $500 000 trucks. On the basis of their applications, five applicants are short-listed. You now need assessments of their driving ability.

Consultant A offers a scheme by which each driver undergoes an eight-week assessment consisting of six separate assessments of knowledge, six separate assessments of driving skills, assessment on the job by a qualified observer for a period of 100 hours, psychological tests to determine temperament and mental stability, and comprehensive health checks for physical impairments.

Total cost $45 000. Validity and reliability of results guaranteed. If a driver assessed as competent has any blameworthy accident in the first year on the job, 100% of fees will be refunded.

Consultant B offers a five hour driving test by a qualified assessor involving all contingencies normally encountered on the job, plus a formal two hour oral assessment of knowledge. All applicants will be asked to produce a medical certificate of fitness to drive from their doctor.

Total cost $2 000. Validity and reliability of results are such that Consultant B can prove that more than 80% of the 400 drivers assessed as competent in the past have not had accidents in their first year on the job.

Whilst it would be reassuring to base your decision on the assessments provided by Consultant A (assessments which obviously have extremely high validity) the cost is prohibitive. Consultant B offers a realistic compromise. Although the assessments do not have an overall validity which matches that of Consultant A, they are sufficiently valid to enable a reasonably safe decision to be made.

BIAS IN ASSESSMENT

The following examples warn against bias in the design and conduct of any assessment.

Content bias

When a revised version of an internationally used test was initially introduced into Australia from USA it contained knowledge questions about US presidents, American geography and the American Civil War. Most Australians, not being as familiar with these topics as their American counterparts probably found these questions more difficult. As a consequence, the test would have produced a result which was slightly more favourable to Americans than Australians. Content bias, such as this example illustrates, detracts from the validity of a test. (The test, a highly reputable one, was subsequently adapted for Australian use by substituting alternative questions.)

Language bias

Some tests of general ability are based on questions which test a candidate's vocabulary and ability to understand concepts and synthesise information, the result being presented as a coherent and logical whole. Because such tests require facility in English language, those for whom English is a second language do not generally do as well. Their results are poorer, not because they lack general ability, but because they lack ability in English. The tests are therefore invalid, due to language bias, as a measure of general ability for such students.

All Australians

Bias due to gender and age

Tests of mechanical reasoning are commonly used in vocational guidance and as a tool for the selection of persons for careers in mechanical engineering fields. These tests measure a person's ability to understand mechanical concepts and apply mechanical principles. Because female students, even in today's enlightened society, still do not receive the same opportunities and encouragement to participate in and pursue mechanically related interests, their mechanical comprehension, as measured by such tests, will be correspondingly lower than that of males.

BEST COPY AVAILABLE

Unless some adjustment is made to take this difference into account such tests are likely to be biased against females. Many tests use tables of norms to compensate for age, as well as gender bias.

Bias against persons with disabilities

Assessments, especially timed tests, can be biased against persons with disabilities. Dyslexic persons, for example, cannot be expected to perform as well in assessments requiring reading or interpretation of other symbolic information. Cerebral palsied persons may have understandable difficulty in tasks requiring fine motor skills. Hearing and sight impairments can also affect assessment performance. A disability does not necessarily mean a person cannot perform a task, only that they may take longer to master it. Competency-based training and assessment can provide such students with a better opportunity to prove their capacity to perform.

Personal prejudice

The effects of bias due to personal prejudice are especially important in assessments based on interviews and oral testing. Because the direction and extent of questioning is to a large degree controlled by the assessor, it is easy, if a candidate is favoured, for an assessor (consciously or otherwise) to allow extra opportunities for responses and to encourage or lead the candidate towards correct answers. To help reduce the possibility of this type of bias it is best to use an interview panel and have all interviewers trained in the concepts and application of the principles of equal opportunity. The effects of personal prejudice can also be

reduced by well structured, objective, oral assessment procedures.

PLANNING ASSESSMENT

There are five important considerations to take into account when planning assessment. These are shown below with the questions you will need to answer.

Authenticity	Is the evidence upon which the assessment is based actually the work of the trainee?
Currency	Does the assessment evidence reflect the current situation or is it out of date?
Cost-effectiveness	Is the assessment method financially realistic?
Validity	Is the assessment measuring what it is supposed to be measuring?
Reliability	Is the assessment consistent in its results?
Fairness	Can all trainees equally demonstrate their competence?

RECOGNISING PRIOR LEARNING

We have become somewhat inflexible in our approach to training and to assessment. We have come to believe that training must take place in a fixed place, at a fixed time, over a fixed period. But does it matter how, where, when and for how long people train? If they are capable of showing that

they are competent to perform a task, then that should be recognised.

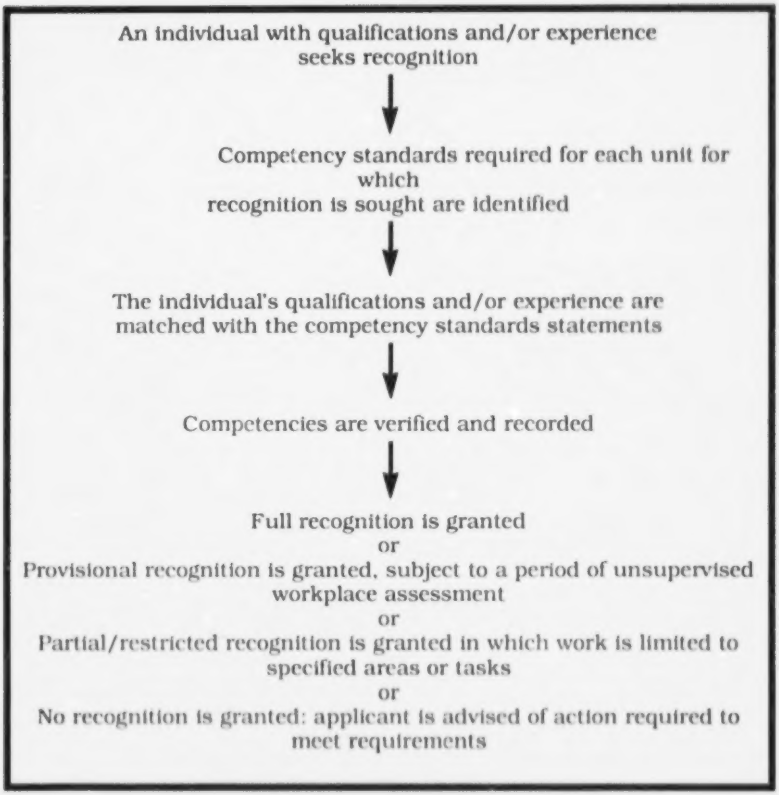
Competence to perform a task may be attained through formal training. It may also be attained informally on the job; or it may be attained informally in other ways. How, when and where learning takes place is not important.

It is now appreciated that prior learning should be recognised through the testing of appropriate skills and knowledge.

Peter Thomson identifies a variety of reasons for seeking recognition of prior learning (RPL). They are:

- to gain entry to a course or training program
- to gain credit in a course or training program
- to gain novice or beginner status in an occupation
- to gain full or competent performer status in an occupation.

He also outlines an approach which may be followed in recognising prior learning. This approach, with minor adaptations, is shown in the following box.


An individual with qualifications and/or experience seeks recognition
Competency standards required for each unit for which recognition is sought are identified
The individual's qualifications and/or experience are matched with the competency standards statements
Competencies are verified and recorded
Full recognition is granted
or
Provisional recognition is granted, subject to a period of unsupervised workplace assessment
or
Partial/restricted recognition is granted in which work is limited to specified areas or tasks
or
No recognition is granted: applicant is advised of action required to meet requirements

The full potential of recognition of prior learning (RPL) will only be realised if:

- explanations and guidelines regarding the RPL process are readily available to all interested parties, particularly applicants and assessors.

- support mechanisms exist in the workplace and education system to encourage and facilitate use of the RPL process by applicants.

ASSESSMENT RESULTS — SOME PROBLEMS AND PITFALLS

This section is a bit technical and may be skipped!

Imagine eight students have each been given a final mark out of 80 which is the average mark for a series of practical projects done during the year and their marks are as shown in column A of the table which follows. There are many ways in which we could treat these scores in order to produce an assessment result. None is absolutely fair. If misused, some can be grossly unfair. The following brief discussion will illustrate a few of the problems and highlight the reasons for the shift to competency-based assessment.

	A	B	C	D	E
Student	**Mark/ 80**	**%**	**Letter Code**	**Percentile Mark**	**[1]CA/CNA**
Judith	72	90	D	93	CA
Leanne	68	85	C	88	CA
Barry	49	61	P1	50	CA
Bev	48	60	P1	48	[2]CNA
Alan	40	50	P2	28	CA
Roger	39	49	F	26	CNA
Tom	28	35	F	9	CNA

1 CA = competency achieved CNA = competency not achieved

2 This student did not achieve competency in one element and was therefore assessed CNA. Note that an assessment which is not competency based might not pick up this lack of competence.

D	=	Distinction	C	=	Credit
F	=	Fail	P1	=	Pass Grade 1
P2	=	Pass Grade			

Column A shows the actual marks (sometimes referred to as 'raw scores') out of 80. Column B shows the marks converted to percentages (e.g. $^{48}/_{80}$ becomes 60%). Column C shows the percentages converted to letter grades.

The actual mark in column A is only of use if the maximum possible score is also known. So, 48 is meaningless unless we are told it is 48 out of a possible 80, which tells us where the person's performance lies in the range of possible scores.

The percentages shown in column B are slightly better because we always know that the maximum score must be 100. At first glance this might suggest that percentages are comparable scores in which case we could say that a student's score of 80% in English is a better performance than that of 60% in Mathematics. We shall see later this is not necessarily true.

Many assessors feel more comfortable with letter grades as shown in column C because they believe it is unrealistic to pretend that assessments can discriminate between students to an accuracy of 1%. They are right. There are just too many sources of variation and error in most assessments. Of course, letter grades do not completely solve this problem: sharp divisions still exist between the grades.

All of us have been assessed in terms of the three types of score described above — most of us have used them ourselves. To illustrate their weakness we need only consider one aspect: the 'pass-mark'. In most of these assessments the passing score is set at half marks or 50%. Why? Because it has always been done that way. Sometimes trainers feel a bit uneasy about this when an unexpectedly large proportion of students passes or fails a test. Some years ago one TAFE college took, what then, was a radical step and increased the pass mark for all tests to 60%. The college did this because the teaching staff felt that the 50% pass mark did not produce a high enough standard. To them at least, the 50% pass mark was not sacred.

This pass mark issue highlights one of the problems of assessment: variations in difficulty between one assessment and another show that we can never be really confident any particular assessment is accurate. We can never be absolutely sure whether the test we are using is too hard, too easy or just right.

One method of overcoming the problem of variations in assessment difficulty is to convert the scores to a norm-

referenced form as in percentile ranks shown in column D. Percentile ranks show where individuals lie with respect to their peers. In simple terms, a percentile rank of 88 means that 87% of the individual's peers did not perform as well and 12% performed better.

Provided the norms used are based on a large number of individuals (called a 'population') our judgement of where an individual lies in relation to others can be fairly accurate. However, we are still no closer to determining a 'correct' pass mark. Should it be at the 50th percentile or should only those above the 75th percentile pass?

The results in column E are based on an entirely different system of assessment and one which overcomes both the problem of variations in test difficulty and that of setting the pass mark. In this type of assessment, often referred to as criterion-referenced or competency-based assessment, the assessment task is defined in terms of a series of learning objectives or elements of a task which must be mastered. For competency to be achieved, each of the objectives or task elements must be successfully completed. The strength of this type of assessment is that it provides an 'absolute' rather than a 'relative' result. We find out whether or not the person can actually perform the task rather than how well they do by comparison with others.

AND LASTLY

For assessment to be effective, it must be part of the educative process, and so

- be based on the achievement of clearly stated course aims (validity)
- be consistent (reliable)
- not be biased against particular groups, i.e. must be fair
- use appropriate techniques
- be a professional matter between the trainee and assessor.

In addition, it must be cost-effective.

FURTHER READING

Ashenden, Dean. 1990, *The recognition of vocational training and learning: A report commissioned by the Employment and Skills Formation Council*, AGPS, Canberra.

This paper considers proposals for advancement in the public accreditation of vocational education and training programs, whether such programs are based in private or public sector institutions. It also deals with ways in which better access can be provided to assessment and certification of learning and competence.

Blank, William E. 1982, *Handbook for Developing Competency Based Training Programs*, Prentice-Hall, Englewood Cliffs, NJ.

The aim of this book is to provide assistance in developing effective education and training programs primarily for instructors, trainers, supervisors and others involved in job-related training. It lists a number of important training program aspects and describes how to develop programs which address them.

Docking, Russell A. 1990, *'A-Z of assessment myths and assessment in the workplace'*, Paper presented at the TAFE National Centre for Research and Development

Conference, "Assessment and standards in vocational education and training" in Adelaide, 1990.

This paper argues that appropriate and rigorous assessment methods can be developed and used in the workplace. Each page deals with a particular assessment *myth* and presents alternative approaches which (at least in theory) can be applied in the workplace.

Employment and Skills Formation Council, 1991, *TAFE in the 1990s: Developing Australia's skills*, National Board of Employment, Education and Training, Canberra.

This paper examines factors shaping TAFE in the 1990s, including TAFE's strengths and weaknesses. New approaches to skills and competence, enterprise level training and TAFE's public obligations are discussed. Long term aims such as increased participation, equity, continuing renewal of curriculum, skill assessment and recognition systems are also considered.

Foyster, John. 1990, *Getting to grips with competency-based training and assessment*, TAFE National Centre for Research and Development, Adelaide.

This booklet is intended to introduce the general reader to the ideas associated with competency-based training and assessment. It defines competence and skills, applies these ideas to two sample programs, and then deals with testing for competency. A final section draws out the implications of this approach to training and assessment.

Foyster, John. 1991, *Getting to grips with on-the-job competencies*, TAFE National Centre for Research and Development, Adelaide.

This booklet deals with training and assessment in the workplace. Ideas to improve training and assessment by using the competency-based approach are examined.

Gifford, Bernard, R & O'Connor, Mary Catherine (eds), 1992, *Changing assessment: Alternative views of aptitude, achievement and instruction*, Kluwer Academic Publishers, Boston.

This book is a collection of articles from selected authors on alternative approaches to testing. The authors were asked to explore new assessment techniques and prepare new measurement tools and methods, such as dynamic assessment, performance-based assessment and portfolio assessment.

Gonczi, Andrew, Hager, Paul & Oliver, Liz. 1990, *Establishing competency-based standards in the professions*, NOOSR research paper no. 1, AGPS, Canberra.

This paper introduces the concept of competency-based standards and illustrates how they may be adapted for professional areas. It considers the concept of standards and discusses the identification of competencies and standards.

Mitchell, Lindsay & Cuthbert, Tommy. 1989, *Insufficient evidence? The final report of the competency testing project*, Scottish Vocational Education Council, Glasgow.

This report is set into two parts. Part one focusses on models of competence, assessment and technical requirements of assessment systems. In part two the results of the project are surveyed and discussed. The report concludes by summarising and examining implications. Descriptive accounts of current practice are included.

National Council for Vocational Qualifications, 1992, *Assessing competence in unpaid work*, NCVQ, London.

This paper considers the acquisition and recognition of competence through unpaid work. Implications for education and training organisations are considered.

National Office of Overseas Skills Recognition. 1991, *Technical qualifications: A summary of guidelines for the assessment of overseas qualifications*, AGPS, Canberra.

This publication contains brief guidelines to assist in assessing overseas qualifications for many technician or para-professional occupations. It considers qualifications from thirty-six countries.

Ness, Stuart, 1989, *Final report of the innovative approaches to assessment project*, SCOTVEC, Glasgow.

This project records and illustrates new approaches to collecting assessment evidence. The project worked with further education colleges in Scotland. Several case studies were used and development groups were formed to review and assess the main issues.

Peddie, Roger, A. 1992, *Beyond the norm: An introduction to standards-based assessment*, New Zealand Qualifications Authority, Wellington.

This booklet is an introduction to assessment focussing on standards-based assessment. It is an introductory guide to assessment of units registered within the New Zealand qualifications framework. The booklet also considers selected issues in assessment and presents examples and problems.

Reid, Neil & Gilmore, Alison. 1987, *Test bias! Test bias?*, in *The Best of SET-Assessment*, Australian Council for Educational Research, Melbourne.

Russell, Tim & North, Denise. 1991, *Competence and assessment of skills in the United Kingdom and Singapore*, Aymever, Armidale, NSW.

This report examines the experiences of the United Kingdom and Singapore as education and training changes to a skills focus. Qualifications, curricula, recruitment and careers are affected as each country approaches the concept differently. This report explains their successes and failures and future plans.

Thomson, Peter. 1989, *Development of procedures for the assessment of experiential learning: Report to the Expert Panel in Technical Qualifications*, Council on Overseas professional Qualifications, Canberra.

This report deals with the possibilities of experiential learning assessment helping with requests for recognition of overseas qualifications.

Thomson, Peter. 1991, *Competency-based training: Some development and assessment issues for policy makers*, TAFE National Centre for Research and Development, Adelaide.

This brief paper argues for the need to understand competence in context. It also deals with such assessment issues as the complex relationship between competence and performance and the need to recognise the problems to be faced in assessment of prior learning.

Thomson, Peter & Pearce, Pat. 1990, *Testing times: A study of assessment on and off the job*, TAFE National Centre for Research and Development, Adelaide.

This report describes the outcome of a study of the assessment procedures which are in place both on and off the job. The study was based on consultations and interviews with individuals in a range of industries in South Australia.

Vocational Education Employment & Training Advisory Committee 1993, *Competency standards — assessors*, Canberra.

This booklet contains cross industry guideline standards to apply to all people who carry out assessment. These standards have been endorsed by the National Training Board and will form a basis for training of those involved with the assessment process.

Whiteley, John. 1991, *CBT: Assessment models appropriate for competency-based training and their relationship to teaching and learning approaches*, Office of Vocational Education, Training and Employment Commission, Brisbane.

This paper is the final report of a national competency-based training pilot project. It reviews research literature on assessment and competency-based training and identifies some broad guidelines regarding good assessment practices within competency-based training.

Wood, Robert et al. 1989, *Boning, blanching and backtracking: Assessing performance in the workplace*, The Training Agency, Sheffield.

This is a technical report of a study of work-based assessment in a variety of settings. The earlier sections (especially section 2) provide a context for thinking about assessment issues as they arise in the workplace, especially such matters as credibility, reliability and

validity. The later sections include 'real life' (and sometimes technical) examples of the application of these ideas.